



Wharton

UNIVERSITY of PENNSYLVANIA

Jacobs Levy Equity  
Management Center  
for Quantitative Financial Research

# Discussion: “Can ChatGPT Forecast Stock Price Movements? Return Predictability and Large Language Models”

Winston Dou (Wharton)

# Motivation

---

## How will AI transform the trading landscape, market efficiency, & stability?

- **Economists:** It touches the heart of asset pricing and capital markets research
- **Practitioners:** It offers opportunities and tools to outperform the market
- **Regulators:** It may raise the risk of market manipulation and instability

## While AI can boost market efficiency, it could have the opposite effect

- **Efficiency enhancements:**
  - Big data capacity
  - Non-structural data processing capacity
  - Optimization capacity for complex and sophisticated dynamic problems, eliminating human errors and emotional biases
- **Efficiency challenges:**
  - AI collusion (Dou\_Goldstein\_Ji, 2024)
  - Herding behavior resulting from homogenization (SEC Report)
  - Compliance difficulties due to the opacity of AI systems (US Congress)
  - Data poisoning for training (US Treasury)

# A difficult question to address

---

## **Question: Is AI a powerful tool or dangerous weapon in financial markets?**

Challenge I: Both aspects exist, which dominates depends on various factors, such as

- The market structure in an asset market
- The concentration of advanced AI technologies
- The ownership of data for training
- The equilibrium that emerges from the interaction of multiple AIs
- The feasibility of regulatory reforms that address AI challenges

Challenge II: Developing a theory for how AI operates or converges is challenging

- Experimental studies on the behavior of AI algo using “synthetic data”  
e.g., Calvano\_Calzolari\_Denicoló\_Pastorello (2020), Dou\_Goldstein\_Ji (2024)
- Empirical studies on the behavior of ML/AI algo using real data  
e.g., Chen\_Kelly\_Xiu (2022), Lopez-Lira\_Tang (2024)

# The objective of this paper

---

**State-of-the-art LLMs, like ChatGPT-4, can extract signals from public data**

- Predict short-term future returns, typically within 1-2 days on average
- Generate substantial daily trading profits through a long-short portfolio strategy

The widespread adoption of LLM-based signals and trading  $\implies$  market efficiency  $\uparrow$

## **Procedure:**

Step 1: Prompt ChatGPT to interpret news headlines as good, bad, or neutral

- Low-frequency, sparse-language-based public information for each stock

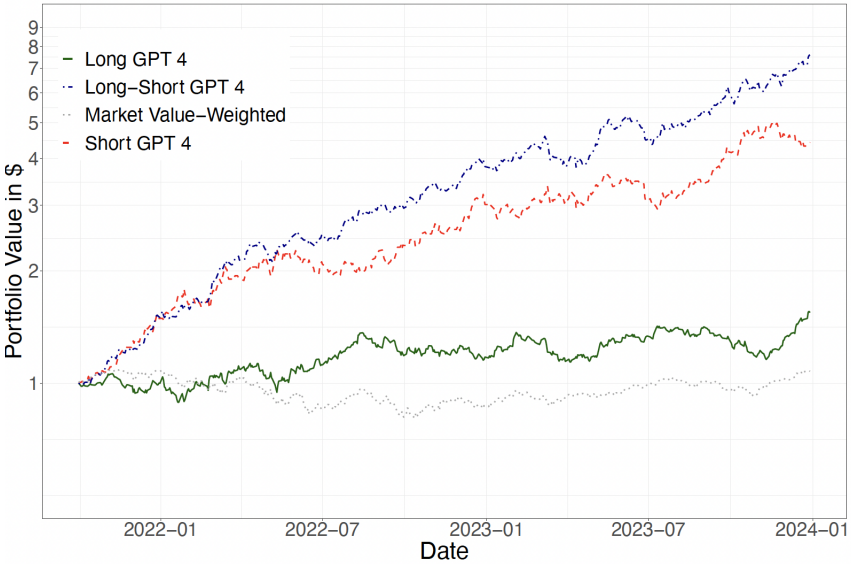
Step 2: Construct long-short trading strategies based on ChatGPT's signals

- Go long on stocks labeled as "good" and short those labeled as "bad"

Step 3: Use linear regression as a surrogate model to understand LLMs

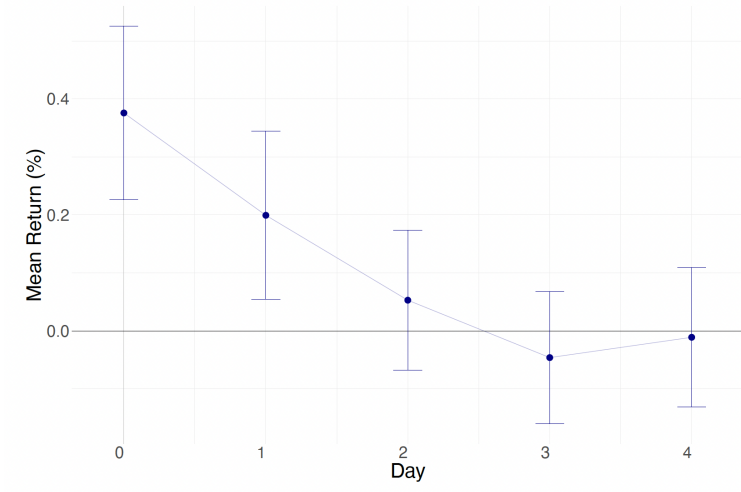
- Regress LLM-generated labels on topic dummy variables of news headlines

# Cumulative return of investing \$1 (no trading costs)



# Returns of overnight news strategy over event time

---



# Heterogeneity across LLMs, news complexity and type

Model	All	Low Complexity	High Complexity	News Articles	Press Releases
GPT-4	<u>3.28</u>	<u>2.60</u>	<u>1.45</u>	<u>2.55</u>	<u>2.10</u>
GPT-3.5	1.79	<u>2.61</u>	<u>0.21</u>	<u>1.92</u>	<u>0.99</u>
DistilBart-MNLI-12-1	1.61	<u>1.53</u>	<u>0.22</u>	<u>1.81</u>	<u>0.49</u>
Ravenpack	1.39	<u>2.17</u>	<u>0.52</u>	<u>2.94</u>	<u>0.82</u>
BART-Large	1.24	1.81	0.45	1.87	1.12
BERT-Large	1.12	-0.29	1.43	0.51	0.75
GPT-1	-0.31	-1.32	0.01	-0.13	0.26
GPT-2	-0.31	-0.45	-0.23	1.17	-0.44
FinBERT	-0.43	-0.66	0.28	-0.30	0.25
BERT	-0.61	-0.17	-0.49	0.54	-0.38
GPT-2-Large	-0.93	-0.30	-1.03	0.08	-0.80

- The complexity score is based purely on sentence length and word complexity
- It doesn't account for the content's meaning, structure, or tone, which can affect overall comprehension

# 1. Economic mechanism

---

## **Here: Slow learning caused by limited information processing capacity**

- It pertains to **public data**, not private information
- Faster learners may strategically trade slowly to collectively maximize their rents
- In this case, LLM-based trading strategies will not enhance market efficiency

## **How to strengthen the economic mechanism?**

- Return predictability and LLM trading profits are especially pronounced for firms in a poor information environment (low **analyst coverage** or low **institutional ownership**)  
e.g., Cohen\_Frazzini (2008), Dou\_Wu (2024)

## **News complexity is a good idea, but what the measure captures is unclear**

- News articles use shorter sentences and simpler words than press releases
- Is it about news complexity or news type?



## 2. Implementable strategies?

---

### **Substantial trading costs include price impact, fees, and margin requirements**

- Under-diversified portfolio with several stocks in both short and long positions, adjusted daily
- Margin calls and funding risk are real concerns for LLM long-short traders
  - Daily volatility  $\approx 2\%$  with daily mean  $\approx 0.4\%$  (Liu\_Longstaff, 2004)

### **Super transitory signals rarely survive high trading costs**

e.g., Jensen\_Kelly\_Malamud\_Pedersen (2024)

- The long-short portfolio must be entirely re-balanced daily  $\Rightarrow$  transaction costs  $\uparrow$

### **Suggestion:** Include trading costs in profits and portfolio optimization

- Consider smart implementable trading strategies

### 3. AI herding behavior driven by “homogenization”

---

#### **Homogenization will emerge endogenously in equilibrium**

- All LLM traders will opt to use the most powerful LLM
- All LLM traders will opt to use the most informative public data
- All LLM traders act on the same signals in the same way

#### **There is no perfect LLM for summarizing news with a single score**

- LLMs will be wrong on some (probably many) topics
- All LLM traders herd in the wrong direction, leading to significant mispricing and market instability

**Suggestion:** Exercise caution when discussing the impact of LLM-powered trading on market efficiency

# Conclusion

---

- **An intriguing paper on a highly important topic**
- **What I appreciate the most:**
  - A useful approach to quantifying information processing capacity
  - A valuable perspective on the impact of AI on market efficiency
- **Suggestions:**
  - Strengthen the economic mechanism
  - Consider implementable strategies accounting for trading costs
  - Discuss potential market inefficiencies and instability risks due to AI